**Инструкция**

Данная программа создает приложение (с помощью библиотеки PyQt5) для рекурсивной распаковки, сортировки файлов по папкам в зависимости от формата а также извлечения данных пользователей из файлов. Программа возвращает excel файл, содержащий в себе следующие столбцы:

**Surname :** Фамилия

**Name :** Имя

**username :** логин (такой же как email)

**email :** email

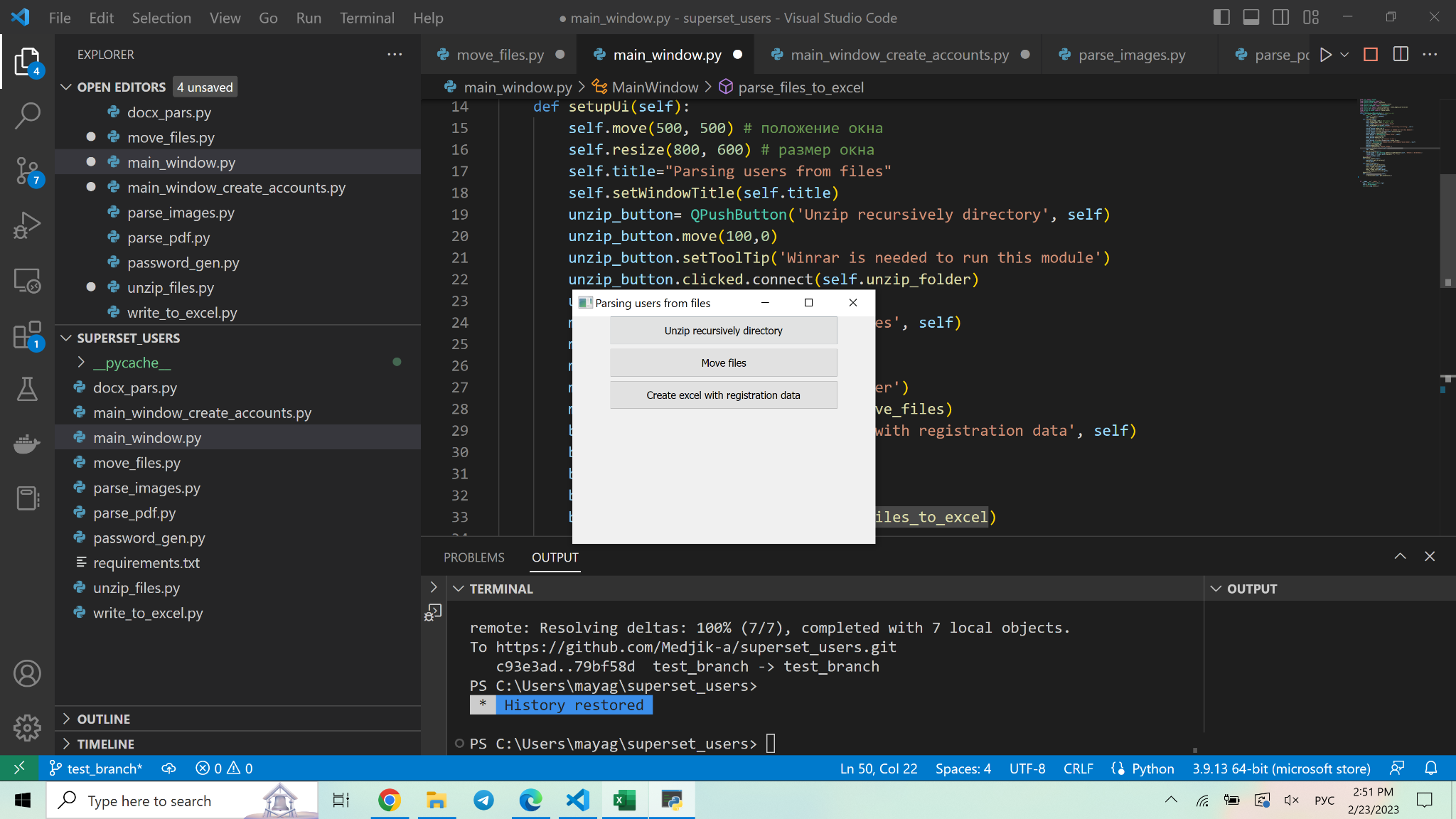
**role :** пользовательская роль в суперсете (Dashboard Viewer)

**password :** сгенерированный пароль

**email\_is\_filled :** удалось ли извлечь email (yes - да, no - нет)

**file :** из какого файла извлечены данные

Чтобы запустить приложение, запустите программу **main\_window.py**. Интерфейс выглядит так:



Чтобы выполнить действие нажмите на нужную кнопку и выберите директорию.

Функции, отвечающие за рекурсивную распаковку находятся в файле **unzip\_files.py.** Функции, сортирующие файлы по формату находятся в файле **move\_file.py**. В файлах **docx\_pars.py**, **parse\_images.py** и **parse\_pdf.py** содержатся функции отвечающие за извлечение информации из docx, изображений и pdf файлов соответственно. В файле **password\_gen.py** содержится функция для генерации сложных паролей. requirements.txt содержит в себе необходимые библиотеки и программы для работы приложения. Для дальнейшей работы с PyQt5, необходимо использовать [следующую документацию](https://doc.qt.io/qtforpython/). Также необходимо загрузить [poppler](https://blog.alivate.com.au/poppler-windows/), [tesseract](https://github.com/UB-Mannheim/tesseract/wiki), [winrar](https://www.win-rar.com/start.html?&L=4) (Инструкции по установке [poppler](https://stackoverflow.com/questions/18381713/how-to-install-poppler-on-windows) и [tesseract](https://linuxhint.com/install-tesseract-windows/)).